UCSF
UC San Francisco Previously Published Works

Title

Amino acid sequence and distribution of mRNA encoding a major skeletal muscle laminin binding protein: an extracellular matrix-associated protein with an unusual COOH-terminal polyaspartate domain.

Permalink

https://escholarship.org/uc/item/6x52s7xq

Journal

The Journal of cell biology, 107(2)

ISSN

0021-9525

Authors

Clegg, DO
Helder, JC
Hann, BC
et al.

Publication Date

1988-08-01

DOI

10.1083/jcb.107.2.699

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Amino Acid Sequence and Distribution of mRNA Encoding a Major Skeletal Muscle Laminin Binding Protein: An Extracellular Matrix-associated Protein with an Unusual COOH-Terminal Polyaspartate Domain

Dennis O. Clegg, Judith C. Helder, Byron C. Hann, Deborah E. Hall, and Louis F. Reichardt

Howard Hughes Medical Institute and Department of Physiology, University of California, San Francisco, California 94143-0724

Abstract. Two cDNAs encoding an abundant chicken muscle extracellular matrix (ECM)-associated laminin-binding protein (LBP) have been isolated and sequenced. The predicted primary amino acid sequence includes a probable signal peptide and a site for N-linked glycosylation, but lacks a hydrophobic segment long enough to span the membrane. The COOH terminus consists of an unusual repeat of 33 consecutive aspartate residues. Comparison with other sequences indicates that this protein is different from previously described LBPs and ECM receptors. RNA blot analysis of LBP gene expression showed that LBP mRNA was abundant in skeletal and heart muscle, but barely detectable in other tissues. Blots of chicken genomic DNA suggest that a single gene encodes this LBP. The amino acid sequence and mRNA distribution are consistent with the biochemical characterization described by Hall and co-workers (Hall, D. E., K. A. Frazer, B. C. Hahn, and L. F. Reichardt. 1988. *J. Cell Biol.* 107:687–697). These analyses indicate that LBP is an abundant ECM-associated muscle protein with an unusually high negative charge that interacts with both membranes and laminin, and has properties of a peripheral, not integral membrane protein. Taken together, our studies show that muscle LBP is a secreted, peripheral membrane protein with an unusual polyaspartate domain. Its laminin and membrane binding properties suggest that it may help mediate muscle cell interactions with the extracellular matrix. We propose the name "aspartactin" for this LBP.

THE basal lamina that surrounds muscle fibers contains molecules that regulate myotube differentiation, promote axon outgrowth, and direct proper synapse formation. One major basal lamina constituent, the glycoprotein laminin, may play roles in all of these functions. It serves as a substrate for myoblast adhesion and migration, and promotes myoblast proliferation and differentiation (Ocalan et al., 1988; Kuhl et al., 1986; Borg et al., 1984; Olwin and Hall, 1985). It is a potent inducer of neurite outgrowth (Manthorpe et al., 1983), and it increases the size and quantity of acetylcholine receptor clusters on primary muscle cells (Vogel et al., 1983; Daniels et al., 1984).

Laminin is a large ($\sim 10^6$ M_r) cruciform glycoprotein complex assembled from three polypeptide subunits. Specific regions of laminin have been shown to interact with other matrix constituents, including collagen IV and heparan sulfate proteoglycans, and with many different cell types (reviewed in Liotta et al., 1986). Three classes of putative cell surface receptors for laminin have been described. These are a "high affinity" M_r 68,000 receptor (Lesot et al., 1983; Terranova et al., 1983; Malinoff and Wicha, 1983; Wewer et al., 1986, 1987), and a lower affinity "integrin" class receptor (Horwitz et al., 1985; Hynes, 1987; Ruoslahti and Pierschbacher, 1987), both of which have been observed on muscle surfaces. Finally, M_r 110,000 and 180,000 laminin-binding proteins have been purified from neuroblastoma cells (Smalheiser and Schwartz, 1987; Kleinman et al., 1988). They do not appear to be present at high levels in skeletal muscle (Kleinman et al., 1988). ECM constituents that bind laminin include entactin (Carlin et al., 1981; Hogan et al., 1982), collagen IV (Terranova et al., 1980), and heparan sulfate proteoglycans (Sakashita et al., 1980; Del Rosso et al., 1981). Molecules that interact with laminin are of considerable interest since they may regulate or mediate laminin actions.

In another report, we describe a novel laminin-binding protein (LBP)[1] that was purified from avian and rat muscle (Hall et al., 1988). This protein is different from previously described proteins and has a strong affinity for membranes

D. O. Clegg's present address is Neuroscience Research Institute and Department of Biological Sciences, University of California, Santa Barbara, CA 93106.

1. *Abbreviations used in this paper*: ECM, extracellular matrix; LBP, laminin-binding protein.

© The Rockefeller University Press, 0021-9525/88/08/699/7 $2.00

as well as laminin. Here we report the molecular cloning of a cDNA for the avian LBP and describe its primary sequence and its mRNA distribution. The LBP sequence is characteristic of a secreted protein and includes an unusual repeat of 33 aspartate residues at its COOH-terminus. Its mRNA is found to be abundant in skeletal muscle and heart, but barely detectable in brain and other organs. Possible roles of the LBP and its poly aspartate tail are discussed.

Materials and Methods

Materials

Restriction enzymes and other enzymes used were purchased from New England Biolabs, Beverly, MA, and Boehringer Mannheim Diagnostics, Inc., Houston, TX. α-[^{32}P]dCTP and γ-[^{32}P]ATP were from Amersham Corp., Arlington Heights, IL. Chickens were supplied by Feather Hill farm, Petaluma, CA.

Isolation of cDNA Clones

We obtained a day-11 chick embryo cDNA library in λgt10 from Drs. Carole Nottenburg and Harold Varmus (Nottenburg and Varmus, 1986). The library contained 900,000 independent isolates before being amplified once by growing on *Escherichia coli* strain C600. We used the mammalian codon usage frequencies compiled by Grantham et al. (1981) to approximate the coding sequence of the NH_2-terminal 20 residues of the avian LBP (Hall et al., 1988). An oligonucleotide 59 bases long (Fig. 2) was synthesized by Genentech Inc. (San Francisco, CA). 11 pmol of the oligonucleotide were end labeled with T4 polynucleotide kinase (Maniatis et al., 1982) to a specific activity of 10^7 cpm/pmol and used to screen a million phage plaques on 20 nitrocellulose filters 150 mm in diameter. Plaques were grown and transferred to filters as described in Maniatis et al. (1982). Filters were prewashed (Maniatis et al., 1982), then prehybridized in 100 ml of 0.1% Ficoll type 400, 0.1% polyvinyl pyrrolidone, 0.1% BSA, 750 mM sodium chloride, 75 mM sodium citrate, 100 μg/ml denatured salmon sperm DNA, 0.1% SDS, 50 mM sodium phosphate, pH 6.8, and 20% formamide at 37°C for 6 h. Probe was added at 3×10^6 cpm/ml and allowed to hybridize for 16 h at 37°C. Filters were then washed in 30 mM sodium chloride, 3 mM sodium citrate, 0.1% SDS at 37°C, and exposed overnight to x-ray film using two Cronex intensifying screens (E. I. Du Pont De Nemours and Co., Wilmington, DE). 10 positive plaques that fell into two classes by restriction enzyme analyses were purified and analyzed (Fig. 1).

DNA Sequencing

Eco RI fragments of two independent clones, λ LNR1 and λ LNR8, were subcloned into the M13 vector M13 mp18 in both orientations. Sequential series of subclones for use in DNA sequencing were prepared using single-stranded DNA and a complementary DNA oligomer to form a specific site for Hind III cleavage (Dale et al., 1985). Linearized single-stranded DNA was digested using the 3′ exonuclease activity of T4 DNA polymerase. Short dA homopolymer tails were added using terminal transferase. The ends of the tailed DNA were joined by annealing with a synthetic DNA oligomer and treating with T4 DNA ligase. Nucleotide sequences were determined using the dideoxynucleotide chain termination method (Sanger et al., 1977). Sequences were edited using the *Gel* sequence management program on Bionet (Intelligenetics Corp., Palo Alto, CA). The sequence for λ LNR1 (1,965 bases) was assembled using information from 93 sequences. The sequence for λ LNR8 (1,315 bases) was assembled using 54 sequences. To determine the sequence at the internal Eco RI site in each clone, DNA fragments containing the intact inserts were purified from Hind III, Bgl II digests of λ LNR1 and λ LNR8. After subcloning into M13 mp18 and M13 mp19, the sequences flanking the Eco RI site were determined using as a primer a synthetic deoxyribonucleotide oligomer corresponding to bases 338–354 in Fig. 2.

RNA and DNA Blots

Total RNA was isolated from various tissues of adult male Leghorn chickens by the guanidine–CsCl method of Chirgwin et al. (1979), and the polyadenylated fraction was purified by two cycles of chromatography on oligo(dT)-cellulose (Aviv and Leder, 1972). Samples of polyadenylated RNA (0.5 and 1 μg) were separated on 1.5% agarose gels containing 2 M formaldehyde (Maniatis et al., 1982) and transferred to Gene Screen nylon membranes (New England Nuclear, Boston, MA) by electroblotting. Transfer was carried out at 4°C in 5 mM Tris, pH 7.8, 2.5 mM sodium acetate, 0.25 mM EDTA for 30 min at 10 V followed by 120 min at 50 V. Blots were UV irradiated (Church and Gilbert, 1984) and prehybridized for 16 h at 50°C in 50% formamide, 50 mM sodium phosphate, pH 7.4, 900 mM sodium chloride, 5 mM EDTA, 0.1% ficoll type 400, 0.1% polyvinyl pyrrolidone, 0.1% BSA, 200 μg/ml denatured salmon sperm DNA, and 5% SDS. A single-stranded DNA probe was synthesized from an M13 subclone template that contained a segment of the coding strand from λ LNR1 (nucleotides 407–1,944 in Fig. 2). The probe was made by annealing a sequencing primer to the template and incubating with the Klenow fragment of DNA polymerase as described by Large et al. (1986). The blot was hybridized to the probe for 24 h at 50°C in a solution with the same composition as the prehybridization mixture. Blots were then washed at 65°C in 0.1% SDS, 15 mM NaCl, 1.5 mM sodium citrate, pH 7. mRNA sizes were determined by comparison to ribosomal RNA bands and denatured DNA markers. For Southern blot analysis (Southern, 1975), genomic DNA was purified from day-8 embryonic chickens (Maniatis et al., 1982), digested with restriction enzymes at 37°C for 16 h, separated on 0.7% agarose gels in 89 mM Tris-borate buffer, pH 8.3, 2 mM EDTA, and then transferred to Gene Screen Plus (New England Nuclear) nylon membranes by capillary transfer. Prehybridization, probe synthesis, hybridization, and washes were carried out as above.

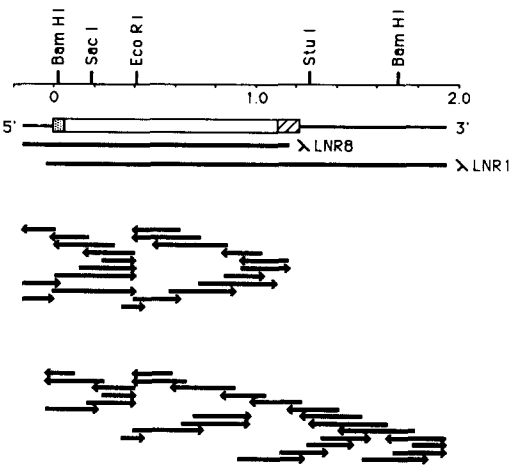

Figure 1. LBP cDNA clones and sequencing strategy. The open box represents the LBP coding sequence, showing the signal peptide (*stippled box*), and the polyaspartate repeat (*hatched box*). Two independent recombinant phage, λLNR8 and λLNR1, are indicated by solid lines, and sequences obtained from their derivative m13 subclones are indicated by arrows beneath each phage. Several restriction enzyme sites are shown on the scale (kilobases) at the top (Bam HI, 18 and 1,698; Sac I, 185; Eco RI, 407; Stu I, 1,261).

Results

Isolation of cDNA Clones

A ^{32}P-end labeled 59-mer oligonucleotide probe, whose sequence was based on the NH_2-terminal amino acid sequence of the avian LBP (Hall et al., 1988) and mammalian codon usage frequencies (Grantham et al., 1981), was used to screen an 11-d chick embryo gt10 cDNA library (Nottenburg and Varmus, 1986). Several recombinant phages that hybridized to the probe were isolated and purified. Restriction enzyme digests indicated that these represented two different classes of cDNAs with different ends derived from the same mRNA (Fig. 1). One member of each class was characterized in detail.

The composite sequence of the LBP mRNA determined

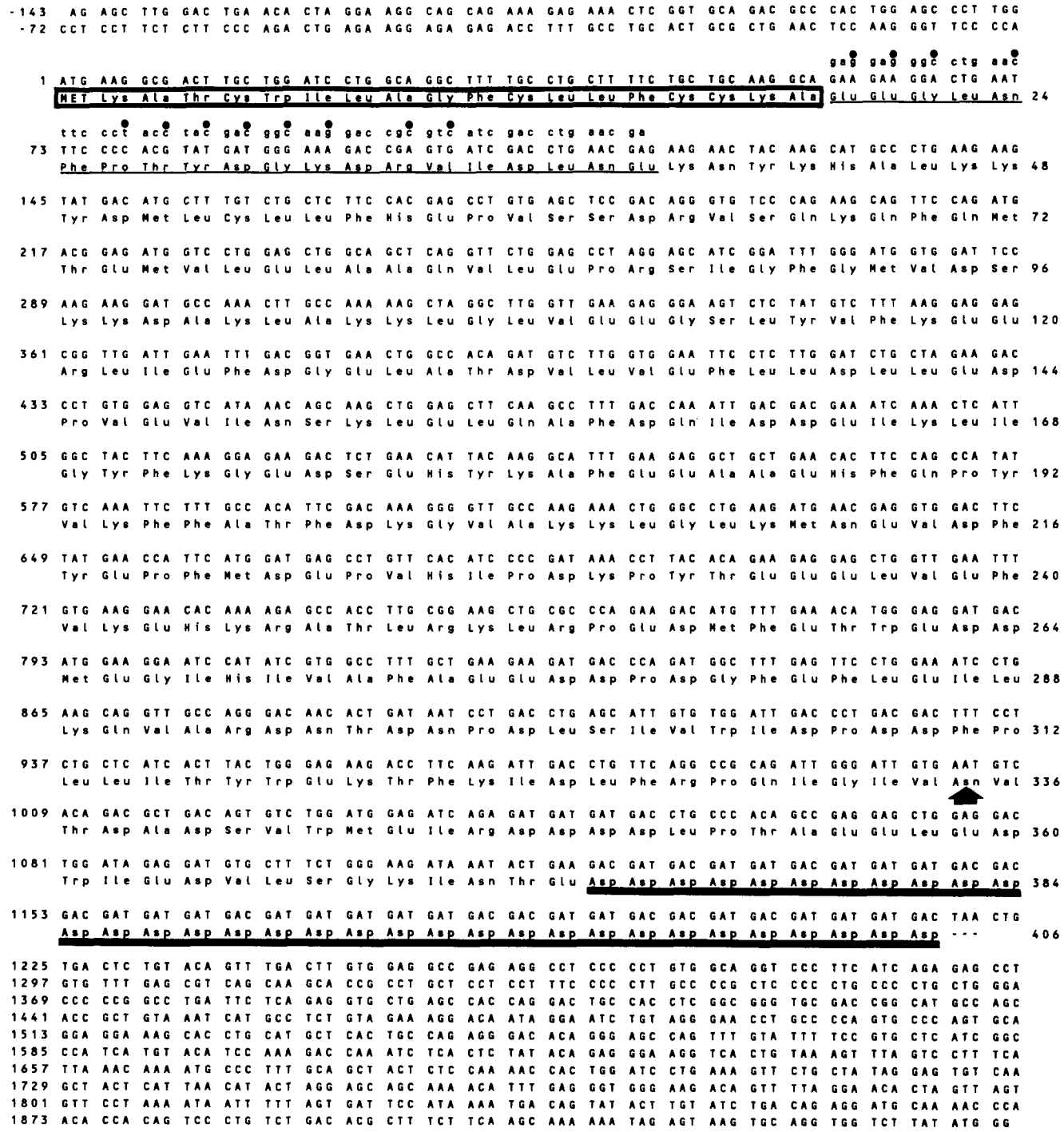

Figure 2. Nucleotide sequence of the LBP cDNA and translated amino acid sequence. Numbers on the left refer to nucleotides and numbers on the right refer to amino acids. The oligonucleotide used to isolate the cDNA clones is indicated in lower case letters above the determined sequence, with mismatches marked by solid circles. The putative signal peptide is boxed, and the NH_2-terminal amino acid sequence derived from the purified protein is underlined. The glycosylation site is denoted by an arrow, and the repeat of aspartate residues at the COOH terminus is underlined with a black bar.

from the two independent cDNA inserts is shown in Fig. 2.[2] The sequence for each clone matches the oligonucleotide probe in 47 of 59 positions, with all of the variation occurring in position 3 of the codons (Fig. 2). The composite sequence is 2,086 bases long, which is approximately the same size as the major hybridizing mRNA in chicken tissues (see below). The two clones overlapped for 1,204 bases and had identical sequences in this region, which includes all of the predicted open reading frame except for the COOH-terminal 46 nucleotides.

2. These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number 400789.

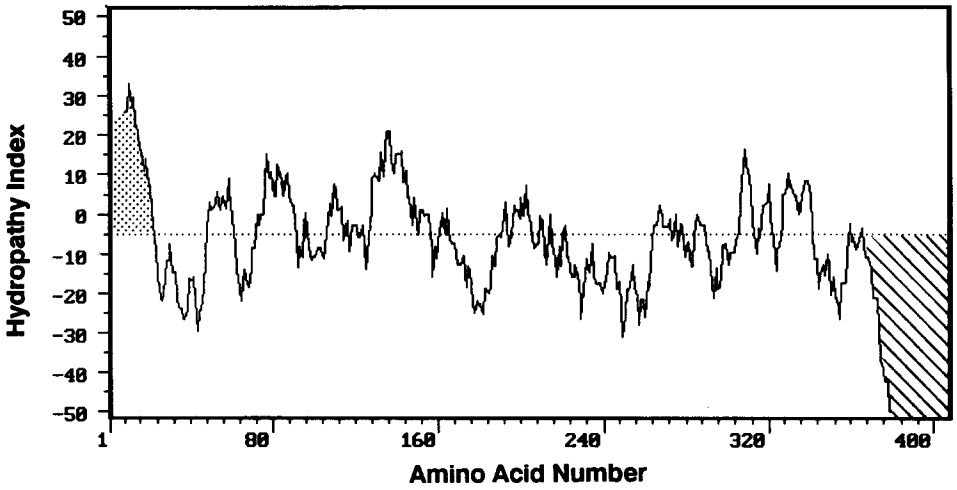

Figure 3. Hydropathy index computation for the LBP amino acid sequence. The sequence was analyzed by the method of Kyte and Doolittle (1982) with the SOAP computer program using a 9 amino acid window of analysis. The hydrophobic regions have a positive hydropathy index score (*left scale*) and appear above the central line. The hydrophobic signal peptide (*stippled area*) and the hydrophilic poly aspartate tail (*striped area*) are indicated.

Features of the LBP Sequence

An open reading frame of 1,218 nucleotides (406 amino acids) was detected that started with an ATG and included sequence encoding the first 20 amino acids determined from the purified protein (Hall et al., 1988). The proposed initiator ATG at position 1 appears to be an exception to the consensus ribosome initiation site (Kozak, 1984, 1986). The ATG is surrounded by a sequence (CCCCAATGA) that matches Kozak's consensus ribosome initiation site (CC(A/G)CCATGG) in only three out of six positions and lacks the highly conserved (95%) purine at −3 (relative to the ATG). As there is an in phase upstream stop codon at position −132 and no other potential upstream initiation codons, the methionine codon at position 1 must serve as the initiation codon.

The predicted amino acid sequence begins with 19 amino acids (boxed in Fig. 2) that have features common to signal peptides. (Perlman and Halvorsen, 1983; von Heijne, 1985, 1986). The putative signal sequence contains a core of 15 hydrophobic amino acids bounded by lysine residues and ends with a small amino acid (alanine). This signal peptide must be removed during normal synthesis to give rise to the observed NH_2-terminal sequence of the mature protein.

In addition to the signal peptide, LBP manifests other properties of a secreted protein. First, there is a single consensus site for N-linked glycosylation (Asn 335) in the COOH-terminal portion of the sequence. This site is likely to be modified since LBP binds to concanavalin A and wheat germ agglutinin–Sepharose columns, and treatment with N-glycosidase F reduces LBP's apparent molecular mass on SDS–polyacrylamide gels by 2,000 D (data not shown). Second, the LBP sequence lacks a hydrophobic sequence of sufficient length to be a probable membrane-spanning domain. The hydrophobicity profile is shown in Fig. 3. Aside from the signal peptide, there are no hydrophobic stretches long enough to span the membrane in an alpha helical conformation, and the computer prediction scheme of Eisenberg et al. (1984) failed to find a suitable membrane spanning region. This was somewhat surprising in view of this protein's strong affinity for both artificial and natural membranes (Hall et al., 1988).

The most striking feature of the LBP primary sequence is the high density of negatively charged amino acids, including an unusual run of 33 aspartate residues at the COOH terminus. In the COOH-terminal half of the protein, 40% of the amino acids are negatively charged, contributing to a net charge of −81 and a predicted isoelectric point of 3.63. The DNA sequence encoding the polyaspartate tail is not a direct repeat but is a mixture of aspartate codons GAT and GAC. The sequence deduced is unlikely to be a cloning artifact because it is included in both independent clones. In addition, results in Table I show that the predicted amino acid composition deduced from the DNA sequence is virtually identical to that determined by analysis of the purified protein. In particular, the predicted aspartate/asparagine content is very close to that observed experimentally.

The mature LBP is 387 amino acids long and has a predicted molecular mass of 45,018 D. This is considerably lower than the 54 kD estimated by migration on SDS–polyacrylamide gels after N-glycosidase F digestion (data not shown), which may be explained by O-linked carbohydrate or may reflect unusual behavior on gels due to the protein's high density of negative charge.

Table I. Comparison of LBP Amino Acid Compositions

Amino acid	Protein analysis	DNA sequence analysis
Asp/Asn	20.0	21.2
Ser	3.2	2.8
Thr	3.8	3.4
Glu/Gln	15.0	14.2
Pro	4.2	4.1
Gly	6.2	3.9
Ala	5.2	4.7
Val	6.3	6.2
Met	1.3	2.3
Ile	7.4	5.4
Leu	8.3	9.6
Tyr	2.2	2.6
Phe	4.2	6.2
His	2.0	1.8
Lys	7.0	7.5
Arg	3.0	2.6

Comparison of LBP amino acid composition determined by protein analysis and DNA sequence analysis. The amino acid composition of the purified protein was determined by hydrolysis and analysis on a Beckman 6300 amino acid analyzer as described by Hall and co-workers. (Hall et al., 1988). The composition was deduced from the DNA sequence using the information in Fig. 2. Both results are expressed as mole percents. Cysteine and tryptophan were not measured on the amino acid analyzer and are not listed.

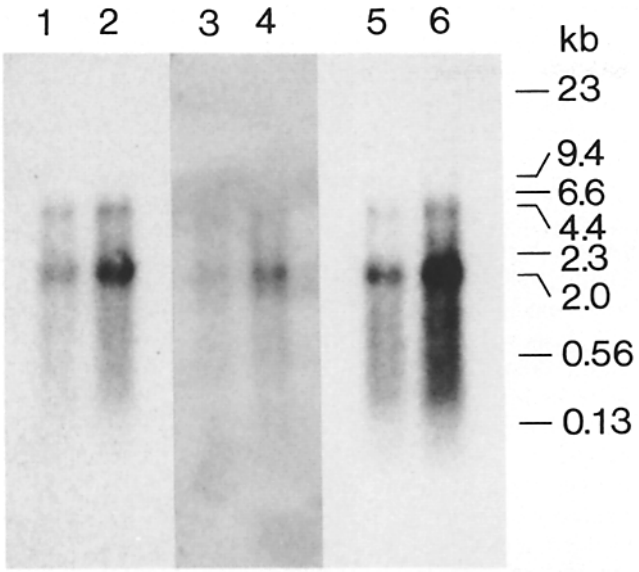

Figure 4. Detection of LBP mRNA in various tissues. The polyadenylated fraction of RNA was hybridized to a LBP cDNA probe (see Materials and Methods). Lanes *1* and *2*, 0.5 and 1 μg heart RNA; lanes *3* and *4*, 0.5 and 1 μg brain RNA; lanes *5* and *6*, 0.5 and 1 μg skeletal muscle RNA. Lanes *3* and *4* were exposed twice as long as the other lanes to show faint bands. The positions and lengths (kilobases) of denatured DNA size markers are shown at right.

Distribution of LBP mRNA

To determine where this LBP is synthesized, polyadenylated mRNA was purified from various tissues of adult chickens and hybridized to a single-stranded DNA probe from the COOH-terminal coding region (see Materials and Methods). A major mRNA band of 2 kb and a minor band of 4 kb were detected (Fig. 4, lanes *1–6*), and their ratio was the same in all tissues examined. The identity of the 4-kb band is not known, but it could be an incompletely processed LBP message or it could be derived from a different, related gene. The 2.0-kb LBP mRNA was most prevalent in skeletal muscle and heart. LBP mRNA could also be detected in brain and kidney, but at levels 50–100-fold below that observed in skeletal muscle (Table II). LBP mRNA was not detected in retina or spleen, liver or testes.

Table II. LBP mRNA Levels

Tissue	Relative amount
Skeletal muscle	100 ± 10
Heart muscle	51 ± 6.5
Brain	2.1 ± 0.3
Kidney	1.3 ± 0.3
Retina, spleen, liver, testes	<0.5

LBP mRNA levels. LBP mRNA was quantified by densitometric scans of the 2.0-kb band detected in various tissues of adult chickens. Two separate preparations of RNA were assayed four different times to yield a mean relative amount ± SEM. Amounts indicated are arbitrary units normalized to the amounts of poly A^+ RNA applied to each gel.

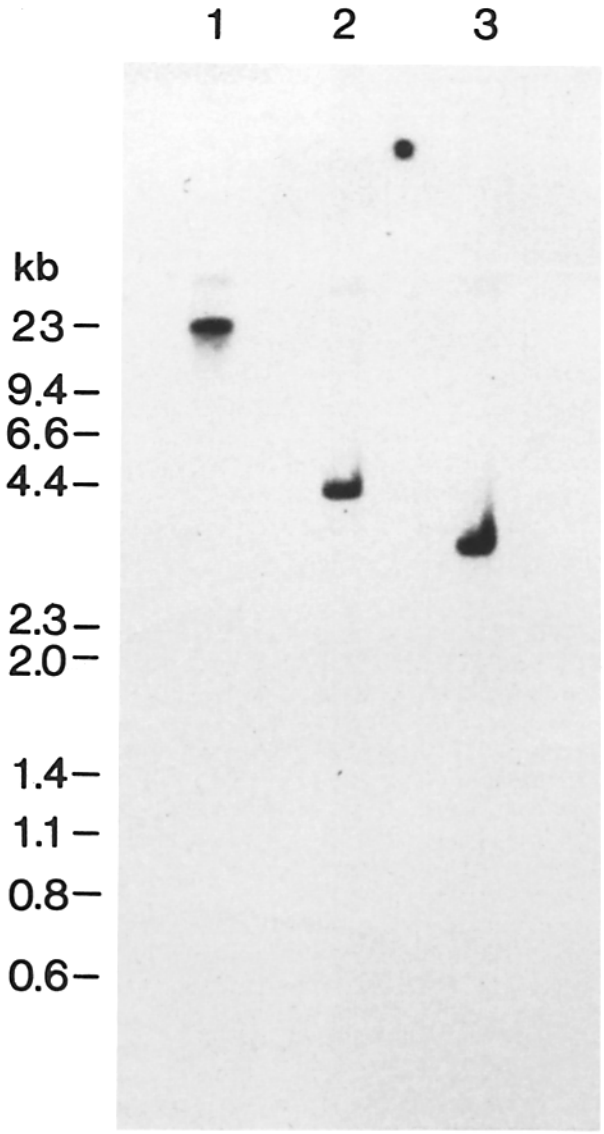

Figure 5. Southern blot analysis of LBP gene. Purified chicken genomic DNA was restricted with Eco RI (lane *1*); Hind III (lane *2*); or Pst I (lane *3*); and hybridized to the LBP cDNA probe (see Materials and Methods). Positions and lengths (kilobases) of size markers are indicated on the left.

Search for Homologous Genes

Southern blot analysis of genomic chicken DNA using a LBP probe revealed a simple pattern of one to two bands (Fig. 5), suggesting that there is only a single LBP gene. We searched the protein sequence database of the National Biomedical Research Foundation for related proteins and found no significant similarities. Recently determined sequences of other proteins thought to interact with laminin, including entactin (Durkin et al., 1987), the beta subunit of the chicken integrin-class ECM receptor family (Tamkun et al., 1986), and the human 68-kD laminin receptor characterized by Wewer et al. (1986, 1987) showed no similarity to the LBP sequence. However, other proteins containing clusters of negatively charged amino acids have been described and some sequences similar to the aspartate repeat were found in the search (see Discussion).

Discussion

The results presented here, coupled with the biochemical characterization made by Hall and co-workers (Hall et al., 1988), describe a novel LBP that differs from those previously reported. Although the molecular mass of the rat LBP is similar to that of the high affinity receptor characterized by others (Terranova et al., 1983; Lesot et al., 1983; Malinoff and Wicha, 1983; Wewer et al., 1986), the sequence of the avian LBP bears no resemblance to the partial receptor sequence reported by Wewer et al. (1986). Furthermore, synthesis of this LBP appears to be restricted to skeletal muscle and heart. Significant levels of its mRNA were not found in other tissues, as would be expected for the high affinity receptor mRNA. The relative molecular mass and tissue distribution of the muscle-associated LBP are different from those of the other described laminin-binding glycoproteins: cranin and the integrins (Smalheiser and Schwartz, 1987; Ruoslahti and Pierschbacher, 1987).

The amino acid sequence of this muscle-associated LBP is compatible with its known biochemical properties and distribution. The protein has been shown to be secreted and localized in the basal lamina surrounding individual muscle fibers (Hall et al., 1988). As predicted by this location, the amino acid sequence of the LBP precursor includes a putative signal sequence that is clearly processed. The mature protein is covalently associated with carbohydrate, as assayed by lectin binding and sensitivity to N-glycosidase F digestion. As predicted, the protein sequence contains a potential site for N-linked glycosylation, which the results of N-glycosidase F digestion imply is actually used as a carbohydrate-acceptor site. While the protein associates with both membranes and phospholipid vesicles, it appears to be a peripheral, not integral membrane protein, since it can be solubilized with high pH and chaotropic agents. Again, this is compatible with the protein's primary amino acid sequence which lacks a hydrophobic sequence long enough to be a potential transmembrane domain. In sections of muscle tissue, LBP immunoreactivity colocalizes with laminin immunoreactivity in the basal lamina. The protein has a strong affinity for laminin, binding to the end of the long arm of the laminin cross (Hall et al., 1988), in the vicinity of the heparin binding site on fragment E3 (Ott et al., 1982). While it is not clear which region of the protein is responsible for binding laminin, there are several acidic regions in the protein that could potentially interact with the heparin binding site.

Sequence analysis revealed a preponderance of negatively charged amino acid residues in the COOH-terminal half of the protein, including an unusual stretch of 33 aspartates at the COOH terminus. In this regard LBP bears some resemblance to the human homeobox protein hox 1.1, which ends with 15 COOH-terminal glutamate residues (Kessel et al., 1987). Repeats of other amino acids at different locations have also been noted in several *Drosophila* homeotic gene products. (Regulski et al., 1985; Wharton et al., 1985*a*). For example, the engrailed gene product includes three near perfect repeats of alanine, glutamine, and glutamate near its NH_2 terminus (Poole et al., 1985). DNA sequences encoding stretches of up to 30 contiguous glutamine residues, known as *opa* or the M repeat, are found several hundred times in the *Drosophila* genome, including in the translated portion of the *Notch* locus (Wharton et al., 1985*b*). Other proteins with equally high densities of negative charge have also been described. For example, ubiquinol-cytochrome *c* reductase from Baker's yeast has 25 contiguous aspartate and glutamate residues (Van Loon et al., 1984). The functions of these repeats in other proteins and in the muscle-associated LBP have not been defined.

The polyaspartate tail of the LBP may explain the extremely high affinity of the protein for hydroxylapatite (Hall et al., 1988), which is known to interact strongly with proteins containing clusters of carboxyl side groups (Gorbunoff, 1984), as well as synthetic polymers of aspartate (Moreno et al., 1984) Nine consecutive aspartic acid residues in the bone ECM protein osteopontin are thought to mediate its tight binding to hydroxylapatate within the bone matrix (Oldberg et al., 1986). The muscle LBP may interact with laminin via its polyaspartate domain at the site(s) on laminin that interact with heparin. Because this muscle LBP is so highly charged, it may also interact with other molecules in the matrix that have clusters of positive charge and contribute to other aspects of basal lamina structure. It should now be possible to determine the role of the polyaspartate tail by constructing and expressing truncated LBP genes.

We can conceive of several general models of LBP function in vivo. In the first model, this protein would serve to anchor muscle cells into the basal lamina by adhering to the surface of the cell membrane and interacting with the long arm of laminin. Alternatively, the muscle LBP could modulate laminin's interaction with other matrix molecules such as heparan sulfate proteoglycans or other negativelv charged competing ligands. It is also possible that this LBP could regulate those cell attachment and neurite outgrowth-promoting activities of laminin which are localized near the LBP-binding site (compare Edgar et al., 1984; Engvall et al., 1986; Hall et al., 1987). Future experiments will be aimed at defining the role served by this protein in the muscle basal lamina.

L. F. Reichardt did much of the molecular biology while on sabbatical in the laboratory of Dr. Stephen Heinemann at the Salk Institute. He thanks Jim Patrick and Steve Heinemann for their hospitality and scientific inspiration. We also thank Dr. J. Ramachandran of Genentech who performed the amino acid composition analysis and NH_2-terminal sequence analysis which made it possible to isolate these cDNAs. At the University of California San Francisco we wish to thank Tom Large for his computer expertise, Terri Burgess for critically reading the manuscript, Brad Olwin and Janet Winter for their early contribution to this project, and Carole Nottenburg and Harold Varmus for the cDNA library.

This research was supported by the Howard Hughes Medical Institute, National Institutes of Health grant NS 19090, and a grant from the Muscular Dystrophy Association. D. O. Clegg was a fellow of the Jane Coffin Childs Memorial Fund for Medical Research and D. E. Hall was a fellow of the Muscular Dystrophy Association. L. F. Reichardt is an investigator of the Howard Hughes Medical Institute.

Received for publication 29 December 1987, and in revised form 6 April 1988.

Note Added in Proof: We propose the name "Aspartactin" for the LBP described here, based on its unusual polyaspartate domain and its probable role in cell-matrix interactions.

References

Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid cellulose. *Proc.*

Natl. Acad. Sci. USA. 69:1408–1412.
Borg, T. K., K. Rubin, E. Lundgren, K. Borg, and B. Obrink. 1984. Recognition of extracellular matrix components by neonatal and adult cardiac myocytes. *Dev. Biol.* 104:86–96.
Carlin, B., R. Jaffe, B. Bender, and A. E. Chung. 1981. Entactin, a novel basal lamina-associated sulfated glycoprotein. *J. Biol. Chem.* 256:5209–5214.
Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active RNA from sources enriched in ribonuclease. *Biochemistry.* 18:5294–5299.
Church, G. M., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA.* 81:1991–1995.
Dale, R. M. K., B. A. McClure, J. P. Houchins. 1985. A rapid single stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. *Plasmid.* 13:31–40.
Daniels, M. P., M. Vigny, P. Sonderegger, H. C. Bauer, and A. Vogel. 1984. Association of laminin and other basement membrane components with regions of high acetylcholine receptor density on cultured myotubes. *Int. J. Dev. Neurosci.* 2:87–99.
Del Rosso, M., R. Cappelletti, M. Viti, S. Vannucchi, and V Chiarugi. 1981. Binding of the basement-membrane glycoprotein laminin to glycosaminoglycans. *Biochem. J.* 199:699–704.
Durkin, M. E., B. E. Carlin, J. Vergnes, B. Bartos, J. Merlie, and A. E. Chung. 1987. Carboxyl-terminal sequence of entactin deduced from a cDNA clone. *Proc. Natl. Acad. Sci. USA.* 84:1570–1574.
Edgar, D., R. Timpl, and H. Thoenen. 1984. The heparin-binding domain of laminin is responsible for its effects on neurite outgrowth and neuronal survival. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1463–1468.
Eisenberg, D., E. Schwarz, M. Komaromy, and R. Wall. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* 179:125–142.
Engvall, E., G. E. Davis, K. Dickerson, E. Ruoslahti, S. Varon, and M. Manthorpe. 1986. Mapping of domains in human laminin using monoclonal antibodies: localization of the neurite-promoting site. *J. Cell Biol.* 103:2457–2465.
Gorbunoff, M. J. 1984. The interactions of proteins with the hydroxyapatite. II. Role of acidic and basic groups. *Anal. Biochem.* 136:433–439.
Grantham, R., C. Gautier, M. Gouy, M. Jacobzone, and R. Mercier. 1981. Codon catalogue usage is a genomic strategy modulated for gene expressivity. *Nucleic Acids Res.* 9:43–74.
Hall, D. E., K. A. Frazer, B. C. Hann, and L. F. Reichardt. 1988. Isolation and characterization of a laminin binding protein from rat and chick muscle. *J. Cell Biol.* 107:687–697.
Hall, D. E., K. M. Neugebauer, and L. F. Reichardt. 1987. Embryonic neural retinal cell response to extracellular matrix proteins: developmental changes and effects of the cell substratum attachment antibody (CSAT). *J. Cell Biol.* 104:623–634.
Hogan, B. L. M., A. Taylor, M. Kurkinen, and J. R. Couchman. 1982. Synthesis and localization of two sulphated glycoproteins associated with basement membranes and the extracellular matrix. *J. Cell Biol.* 95:197–204.
Horwitz, A., K. Duggan, R. Greggs, C. Decker, and C. Buck. 1985. The cell substrate attachment (CSAT) antigen has properties of a receptor for laminin and fibronectin. *J. Cell Biol.* 101:2134–2144.
Hynes, R. O. 1987. Integrins: a family of cell surface receptors. *Cell.* 48: 459–554.
Kessel, M., F. Schulze, M. Fibi, and P. Gruss. 1987. Primary structure and nuclear localization of a murine homeodomain protein. *Proc. Natl. Acad. Sci. USA.* 84:5306–5310.
Kleinman, H. K., R. L. Ogle, F. B. Cannon, C. D. Little, T. M. Sweeny, and L. Luckenbill-Edds. 1988. Laminin receptors for neurite formation. *Proc. Natl. Acad. Sci. USA.* 85:1282–1286.
Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 12:857–872.
Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiation codon that modulates translation by eukaryotic ribosomes. *Cell.* 44:283–292.
Kuhl, U., M. Ocalan, R. Timpl, and K. von der Mark, 1986. Role of laminin and fibronectin in selecting myogenic versus fibrogenic cells from skeletal muscle cells *in vitro*. *Dev. Biol.* 117:628–35.
Kyte, J., and R. F. Doolittle. 1982. A simple method for describing the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132.
Large, T. H., S. B. Bodary, D. O. Clegg, G. Weskamp, U. Otten, and L. F. Reichardt. 1986. Nerve growth factor gene expression in the developing rat brain. *Science (Wash. DC).* 234:352–355.
Lesot, H., U. Kuhl, and K. von der Mark. 1983. Isolation of a laminin-binding protein from muscle cell membranes. *EMBO (Eur. Mol. Biol. Organ.) J.* 2:861–865.
Liotta, L. A., C. N. Rao, and U. M. Wewer. 1986. Biochemical interactions of tumor cells with the basement membrane. *Annu. Rev. Biochem.* 55:1037–1057.
Malinoff, H. L., and M. S. Wicha. 1983. Isolation of a cell surface receptor protein for laminin from murine fibrosarcoma cells. *J. Cell Biol.* 96:1475–1479.
Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
Manthorpe, M., E. Engvall, E. Ruoslahti, F. M. Longo, G. E. Davis, and S. Varon. 1983. Laminin promotes neuritic regeneration from cultured peripheral and central neurons. *J. Cell Biol.* 97:1882–1890.
Moreno, E. C., M. Kresak, and D. I. Hay. 1984. Adsorption of molecules of biological interest onto hydroxyapatite. *Calcif. Tissue Int.* 36:48–59.
Nottenburg, C., and H. E. Varmus. 1986. Features of the chicken c-myc gene that influence the structure of c-myc RNA in normal cells and bursal lymphomas. *Mol. Cell Biol.* 6:2800–2806.
Ocalan, M., S. L. Goodman, U. Kuhl, S. D. Hauschka, and K. von der Mark. 1988. Laminin alters cell shape and stimulates motility and proliferation of murine skeletal myoblasts. *Dev. Biol.* 125:158–167.
Oldberg, A., A. Franzen, and D. Heingard. 1986. Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reveals an Arg-Gly-Asp cell binding sequence. *Proc. Natl. Acad. Sci. USA.* 83:8819–8823.
Olwin, B. B., and Z. W. Hall. 1985. Developmental regulation of laminin accumulation in the extracellular matrix of a mouse muscle cell line. *Dev. Biol.* 112:359–367.
Ott, V., E. Odermatt, J. Engel, H. Furthmayr, and R. Timpl. 1982. Protease resistance and conformation of laminin. *Eur. J. Biochem.* 23:63–72.
Perlman, D., and H. O. Halvorson. 1983. A putative signal peptidase recognition site and sequence in eukaryotic and procaryotic signal peptides. *J. Mol. Biol.* 167:391–409.
Poole, P. J. W., L. M. Kauvar, B. Drees, and T. Kornberg. 1985. The engrailed locus of Drosophila: structure analysis of an embryonic transcript. *Cell.* 40:37–43.
Regulski, M., K. Harding, R. Kostriken, F. Karch, M. Levine, and W. McGinnis. 1985. Homeo box genes of the Antennapedia and Bithorax complexes of Drosophila. *Cell.* 43:71–80.
Ruoslahti, E., and M. D. Pierschbacher. 1987. New perspectives in cell adhesion: RGD and integrins. *Science (Wash. DC).* 238:491–497.
Sakashita, S. E., E. Engvall, and E. Ruoslahti. 1980. The basement membrane protein laminin binds to heparin. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 116:243–246.
Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–467.
Smalheiser, N. R., and N. B. Schwartz. 1987. Cranin: A laminin-binding protein of cell membranes. *Proc. Natl. Acad. Sci. USA.* 84:6457–6461.
Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503–517.
Tamkun, J. W., D. W. DeSimone, D. Fonda, R. S. Patel, C. Buck, A. F. Horwitz, and R. O. Hynes. 1986. Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell.* 46:271–282.
Terranova, V. P., C. N. Rao, T. Kalebic, I. M. Margulies, and L. A. Liotta. 1983. Laminin receptor on human breast carcinoma cells. *Proc. Natl. Acad. Sci. USA.* 80:444–448.
Terranova, V. P., D. H. Rohrbach, and G. R. Martin. 1980. Role of laminin in the attachment of PAM212 (epithelial) cells to basement membrane collagen. *Cell.* 22:719–728.
Van Loon, A. P. G. M., R. J. De Groot, M. DeHaan, A. Dekker, and L. A. Grivell. 1984. The DNA sequence of the nuclear gene coding for the 17 kd subunit VI of the yeast ubiquinol-cytochrome c reductase: a protein with an extremely high content of acidic amino acids. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1039–1043.
Vogel, Z., C. N. Christian, M. Vigny, H. C. Bauer, P. Sonderegger, and M. P. Daniels. 1983. Laminin induces acetylcholine receptor aggregation on cultured myotubes and enhances the receptor aggregation activity of a neuronal factor. *J. Neurosci.* 3:1058–1068.
von Heijne, G. 1985. Signal sequences: the limits of variation. *J. Mol. Biol.* 184:99–105.
von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683–4690.
Wewer, U. M., L. A. Liotta, M. Jaye, G. A. Ricca, W. N. Drohan, A. P. Claywmith, C. N. Rao, P. Wirth, J. E. Coligan, R. Albrechtsen, M. Murdy, and M. E. Sobel. 1986. Altered levels of laminin receptor mRNA in various human carcinoma cells that have different abilities to bind laminin. *Proc. Natl. Acad. Sci. USA.* 83:7137–7147.
Wewer, U. M., G. Taraboletti, M. E. Sobel, R. Albrechtsen, and L. A. Liotta. 1987. Role of laminin receptor in tumor cell migration. *Cancer Res.* 47: 5691–5698.
Wharton, K. A., K. M. Johansen, T. Xu, and S. Artavanis-Tsakonas. 1985*a*. Nucleotide sequence from the neurogenic locus notch implies a gene product that shares homology with proteins containing EGF-like repeats. *Cell.* 43: 567–581.
Wharton, K. A., B. Yedvobnick, V. G. Finnerty, and S. Artavanis-Tsakonas. 1985*b*. opa: a novel family of transcribed repeats shared by the notch locus and other developmentally regulated loci in D. melanogaster. *Cell.* 40: 55–62.